

# Uncertainty for calculating transport on Titan: a probabilistic description of bimolecular diffusion parameters

S. Plessis

Institute for Computational Engineering and Sciences, UT Austin  
`splessis@ices.utexas.edu`

D. McDougall

Institute for Computational Engineering and Sciences, UT Austin

K. Mandt

Space Science Department, Southwest Research Institute  
`kmandt@swri.edu`

T. Greathouse

Space Science Department, Southwest Research Institute

A. Luspay-Kuti

Space Science Department, Southwest Research Institute

May 23, 2017

## Abstract

Bimolecular diffusion coefficients are important parameters used by atmospheric models to calculate altitude profiles of minor constituents in an atmosphere. Unfortunately, laboratory measurements of these coefficients were never conducted at temperature conditions relevant to the atmosphere of Titan. Here we conduct a detailed uncertainty analysis of the bimolecular diffusion coefficient parameters as applied to Titan's upper atmosphere to provide a better understanding of the impact of uncertainty for this parameter on models. Because temperature and pressure conditions are much lower than the laboratory conditions in which bimolecular diffusion parameters were measured, we apply a Bayesian framework, a problem-agnostic framework, to determine parameter estimates and associated uncertainties. We solve the Bayesian calibration problem using the open-source **QUESO** library which also performs a propagation of uncertainties in the calibrated parameters to temperature and pressure conditions observed in Titan's upper atmosphere. Our results show that, after propagating uncertainty through the Massman model, the uncertainty in molecular diffusion is highly correlated to temperature

and we observe no noticeable correlation with pressure. We propagate the calibrated molecular diffusion estimate and associated uncertainty to obtain an estimate with uncertainty due to bimolecular diffusion for the methane molar fraction as a function of altitude. Results show that the uncertainty in methane abundance due to molecular diffusion is in general small compared to eddy diffusion and the chemical kinetics description. However, methane abundance is most sensitive to uncertainty in molecular diffusion above 1200 km where the errors are nontrivial and could have important implications for scientific research based on diffusion models in this altitude range.

## 1 Introduction

Titan atmospheric models contain two major modules: a chemical kinetics module and a transport module. While there have been numerous efforts to quantify uncertainty in the chemical kinetics module [13, 14, 3, 5, 2, 9, 15, 23, 22, 11], the transport module is not as thoroughly evaluated. Furthermore, the transport module consists of two components: an eddy diffusion component and a molecular diffusion component. [4] assessed the question of uncertainty in chemical transport, but only focussed on the eddy diffusion profile effects, since it is the main contributor to uncertainty in the lower atmosphere, [34, Sec. 2.2.3] determined by a Monte Carlo approach a best fit of eddy diffusion profile. However, molecular diffusion plays a critical role in determining altitude profiles in the upper atmosphere and has yet to be investigated cognizant of uncertainty and sensitivity. Parameters for molecular diffusion—molecular binary diffusion coefficients for each pair of molecules—have not been revisited since their initial measurements made in 1973 [32].

[35] reviewed binary diffusion, describing three different sources for the binary molecular diffusion coefficients, all of which are the result of a nonlinear least-squares fit to experimental measurements. These measurements were performed under temperatures ranging from 300 K to 700 K at 1 atm (see Fig. 1). In order to model bimolecular diffusion in Titan’s atmosphere, measurements must be extrapolated to conditions relevant to Titan—approximately 150 K and  $10^{-6}$  atm. Here we evaluate the potential error involved in propagating to temperature and pressure relevant to Titan’s atmosphere by also propagating its uncertainty through this extrapolation.

In this work we apply a probabilistic calibration framework based on Bayes’ theorem. The application of this Bayesian framework solves the problem of estimating, with uncertainty, unknown molecular diffusion parameters in a model for Titan’s atmosphere. We review a brief history of the evolution of bimolecular diffusion models in section 2, and in section 3 we set up the Bayesian framework in which we estimate parameters for bimolecular diffusion. In section 4 we detail the measurements and measurement error of bimolecular diffusion available in the scientific literature and we couple these measurements with a Bayesian calibration procedure to estimate parameters for bimolecular diffusion. This estimate is then propagated through a bimolecular diffusion model to obtain an

estimate of bimolecular diffusion of the  $\text{N}_2/\text{CH}_4$  pair in physical conditions consistent with those typically found on Titan’s atmosphere. Using this, we then estimate the mixture diffusion coefficient of  $\text{N}_2$ . We execute this entire process while taking into account all available experimental and a priori uncertainty. Section 5 concludes and summarises the paper.

## 2 Molecular diffusion models

Molecular diffusion in Titan’s atmosphere is modeled using the Wilke equation [33, 35], or the mixture-average rule, from the bimolecular diffusion coefficient:

$$D_s = \frac{n_{\text{total}} - n_s}{\sum_{j_m \neq s} \frac{n_{j_m}}{D_{s,j_m}}}, \quad (1)$$

where  $n_{\text{total}}$  and  $n_s$  are the total molar concentration and molar concentration of species  $s$ , respectively, and  $D_{s,j_m}$  is the bimolecular diffusion coefficient between species  $s$  and species  $j_m$ . The  $j_m$  notation is used to take into account a possible approximation of the atmosphere’s composition where only the dominant species are considered, thus defining a “medium” in which the molecules are diffusing. In Titan’s atmosphere, this medium is typically composed of  $\text{N}_2$ ,  $\text{CH}_4$  and sometimes  $\text{H}_2$ .

[8] derived a modified version of the mixture-averaged rule applicable for gas mixtures with more than two components. The purpose of this derivation was to address the rapidly changing composition of Titan’s upper atmosphere above the homopause, where binary diffusion is dominant. It required a rederivation of the transport term due to molecular diffusion, adapted to minor species diffusing through a bulk gas,

$$\omega_s^D = -D_s \left( \frac{1}{n_s} \frac{\partial n_s}{\partial z} + \frac{1}{H_s} + (1 + \alpha_T) \frac{1}{T} \frac{\partial T}{\partial z} \right), \quad (2)$$

to a modified version given by

$$\omega_s^D = -\tilde{D}_s \left( \frac{1}{n_s} \frac{\partial n_s}{\partial z} + \frac{1}{H_s} + (1 + \frac{n_{\text{total}} - n_s}{n_{\text{total}}} \alpha_T) \frac{1}{T} \frac{\partial T}{\partial z} \right), \quad (3)$$

where  $H_s$  is the scale height of species  $s$ ,  $\alpha_T$  is the thermal coefficient,  $n_{\text{total}}$  is the total density of the atmosphere and

$$\tilde{D}_s = \frac{D_s}{1 - \frac{n_s}{n_{\text{total}}} (1 - \frac{m_s}{m_{\neq s}})}, \quad (4)$$

where  $m_s$  is the molecular mass of species  $s$  and  $m_{\neq s}$  is the mean molecular mass of the atmosphere without species  $s$ . This expression of molecular diffusion is suited for both minor and major species, with Eq. 3 converging towards Eq. 2 for the minor species. The detailed analysis is given in [8, Section 3.3.1].

Three different models have been proposed in the literature for bimolecular diffusion between any pair of molecules, denoted as the pair of molecules  $s$  and

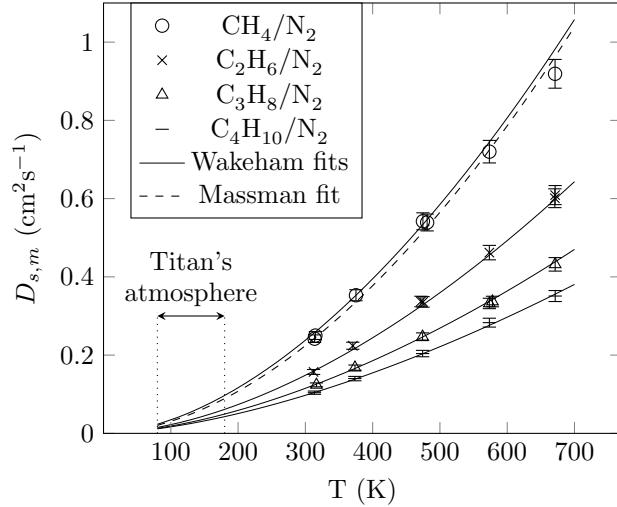


Figure 1: Measurements and fits. The uncertainty in the laboratory measurements is 4%. Titan’s atmospheric conditions are highlighted to emphasize the extrapolation required to apply these values to modeling Titan’s atmosphere.

$m$ . Note that a binary diffusion coefficient is specific to a pair of molecules and is agnostic to which species is diffusing and which species is the medium. First, [32] proposed a model depending only on temperature (Eq. 5),

$$D_{s,m}(T) = AT^s, \quad (5)$$

where  $T$  is the temperature and  $A$  and  $s$  are the model parameters. Note that the temperature should be a reduced unitless temperature: it is more appropriate to write  $\frac{T}{T_{\text{ref}}}$  instead of  $T$  with  $T_{\text{ref}} = 1$  K. The second model, by [35], introduced a pressure dependence,

$$D_{s,m}(T, P) = \frac{A}{n_{\text{tot}}} T^{\beta}, \quad (6)$$

where  $n_{\text{tot}}$  is the total molar concentration,  $A$  and  $\beta$  are the parameters. The pressure dependence is given by the mixture equation of state which, in the case of Titan’s atmosphere, is the ideal gas law,  $n_{\text{tot}} = P/RT$ . Note that, just as above, it is more appropriate to write this with respect to a reference temperature. Finally, [20] also proposed a model with explicit temperature and pressure dependence, but included the bimolecular diffusion,  $D^{(0)}$ , at 0°C and 1 atm ( $T^0$  and  $P^0$  respectively) as a reference value,

$$D_{s,m}(T, P) = D^{(0)} \frac{P^0}{P} \left( \frac{T}{T^0} \right)^{\beta}. \quad (7)$$

with  $D^{(0)}$  and  $\beta$  being the parameters of the model.

Modeling Titan’s atmosphere requires a pressure extrapolation, it is therefore more appropriate to consider either the model of [20] or [35]. Those models are equivalent through the state equation. From a modeler’s point-of-view, the model of [20] has the advantage to extrapolate a reference value of the diffusion coefficient, which holds therefore the potential to be informed independently.

The rest of this work uses the model of [20].

### 3 Bayesian framework

In many scientific applications, some quantities need to be determined from experimental observations that were subject to errors during the experiment. Since the quantity is an extrapolation of measurements, an exact value of this quantity cannot be known with certainty. It is therefore important to provide a measure of the associated uncertainty. We do this for the bimolecular diffusion model described above in a Bayesian setting. The Bayesian setting is advantageous because it provides a *distribution* of a quantity of interest, and this is achieved in two steps. The first step is calibration, where the model is fit to the data to obtain distributions for model parameters. The second step propagates these distributions to a quantity of interest. The Bayesian approach is not a statistical tool, but rather a method for describing an inverse problem. It differs from the simple Monte Carlo technique commonly used in planetary science applications. In the simple Monte Carlo technique one would use a Monte Carlo model as a tool to cast random experimental measurements from the measured mean values and errors, and then apply a Massman model fitting to each set of randomly determined measurements in order to obtain a probability distribution of  $D^{(0)}$  or  $\beta$ . In the Bayesian approach, we use a Markov chain Monte Carlo (MCMC) algorithm (a statistical tool) to solve a Bayesian inverse problem. This approach is very similar to the simple Monte Carlo technique except that the MCMC does not cast random observations, as described above, but  $D^{(0)}$  and  $\beta$ . These proposed values are then compared to the observations and the Metropolis-Hastings acceptance probability decides whether or not the proposed  $(D^{(0)}, \beta)$  sample should be kept. The samples of  $D^{(0)}$  and  $\beta$  at the end of the procedure are realizations from the probability distribution  $p(D^{(0)}, \beta|y)$ , which is desired output for our study.

For the calibration step, we need to describe the model to calibrate and the data to calibrate against. We will calibrate the Massman model for bimolecular diffusion (Eq. 7) against direct noisy observations of  $D_{s,m}$  at various temperatures and pressures,

$$y_{jk} = D_{s,m}(T_j, P_k) + \eta_{jk}, \quad \eta_{jk} \sim \mathcal{N}(0, \sigma_{jk}), \quad (8)$$

where  $j = 1, \dots, J$ , and  $k = 1, \dots, K$ . For the sake of simplicity, we write the observations as a column vector,  $y = (y_{11}, \dots, y_{JK})^\top$ . Here  $\mathcal{N}(0, \sigma^2)$  is notation for a Gaussian distribution with mean zero and variance  $\sigma^2$ . The  $\sigma_{jk}$  will appear later in the expression for the likelihood (see Section 3.2).

The calibration step proceeds by finding the ‘best’ model parameters,  $D^{(0)}$  and  $\beta$ , given the observations  $y$ . That is, we seek the joint probability distribution  $p(\beta, D^{(0)}|y)$ . Applying Bayes’s theorem yields,

$$\underbrace{p(\beta, D^{(0)}|y)}_{\text{posterior}} \propto \underbrace{p(y|\beta, D^{(0)})}_{\text{likelihood}} \underbrace{p(\beta, D^{(0)})}_{\text{prior}}. \quad (9)$$

The prior distribution encodes a previously held state of knowledge, or an external expert opinion, about what the parameters  $\beta$  and  $D^{(0)}$  should look like. It is therefore given. The likelihood distribution is also known; given  $\beta$  and  $D^{(0)}$ , one inserts them into the Massman model and evaluates Eq. 8. Therefore, since the right-hand side of Eq. 9 is known, the left-hand side is also known. The posterior distribution is the solution to the Bayesian calibration problem and there are many numerical methods to understand its properties. We choose to approximate the posterior distribution by statistical samples (see Section 3.3). Although we have applied the Bayesian calibration framework to a specific model, in general the framework is problem-agnostic and can be used in any scientific domain for parameter estimation and uncertainty quantification.

The second step is to propagate the posterior distribution above, which we choose to approximate by samples, to the quantity of interest. For our purposes, the quantity of interest is the bimolecular diffusion  $D_{s,m}(\tilde{T}, \tilde{P})$  at some desired temperature  $\tilde{T}$  and pressure  $\tilde{P}$ . This is very easily done by evaluating

$$D_n^{(0)} \frac{P^0}{\tilde{P}} \left( \frac{\tilde{T}}{T^0} \right)^{\beta_n} \quad (10)$$

for each sample  $n = 1, \dots, N$ . This will yield another sequence of samples of bimolecular diffusions at the desired temperature and pressure. An estimate of the bimolecular diffusion at temperature  $\tilde{T}$  and pressure  $\tilde{P}$  can be obtained by computing the sample mean. The associated uncertainty of this estimate is obtained by computing the sample variance.

### 3.1 Choice of prior

The choice of a prior distribution is a very important part of the Bayesian framework, as it has a direct influence on the posterior’s description (see Eq. 9). In general, this choice can be very difficult. It is a statement about knowledge of the solution of the problem, which is unknown. However there are situations in which one can provide a prior distribution without knowing the solution exactly, but knowing some *property* of the solution. For example, an external observer can never know the exact speed of a car but it is assuredly non-negative and less than the speed of light. Although contrived, this example illustrates that common sense can be harnessed when expert guidance is absent.

When a domain expert is present, one may attempt to acquire high-quality information which can be incorporated into a prior. Of course, the aim here is to do this as accurately as possible. This is still an active area of research for

which no gold standard exists [17]. The process of obtaining expert opinion for a particular problem is called *elicitation*.

When one truly possesses no problem insight, the prior typically relies on the maximum entropy principle (MEP) [29, 10]. This principle states that, given a set of constraints, the distribution that encodes the least amount of information is one which maximizes the Shannon entropy [30]. For instance, if a mean and standard deviation of a parameter are given as constraints, the distribution that maximizes the Shannon entropy is a Gaussian. This method provides a distribution that attempts to maximize ignorance about the solution of the problem. Of course, when a domain expert is available one can still utilise the maximum entropy principle by setting the constraints through expert elicitation.

The parameter  $D^{(0)}$  is a diffusion coefficient and must be positive whilst  $\beta$  is an exponent and therefore can take any real value. With no knowledge other than bounds, the probability distribution that maximises ignorance in the Shannon entropy sense is a uniform distribution. Therefore, a priori, we choose  $D^{(0)} \sim U[0, \infty]$  and  $\beta \sim U[-\infty, \infty]$  so the prior distribution is a joint uniform on the  $(D^{(0)}, \beta)$  pair.

### 3.2 Likelihood

In Eq. 9 we must evaluate the likelihood of observing the data  $y$  given values of the parameters  $D^{(0)}$  and  $\beta$ . Since the experimental data was taken at a fixed set of temperatures  $\mathbf{T} \in \mathbb{R}^J$  and pressures  $\mathbf{P} \in \mathbb{R}^K$ , we define the function  $\mathcal{G}$  to be,

$$\begin{aligned}\mathcal{G}(D^{(0)}, \beta) &= D_{s,m}(\mathbf{T}, \mathbf{P}) \\ &= D^{(0)} \frac{\mathbf{P}^0}{\mathbf{P}} \left( \frac{\mathbf{T}}{\mathbf{T}^0} \right)^\beta,\end{aligned}$$

where division, multiplication and exponentiation are all done component-wise on the vectors  $\mathbf{T}$  and  $\mathbf{P}$ . We then derive the likelihood by recalling from Eq. 8 that the difference of the model output and the observation is a Gaussian random variable. Therefore, the likelihood distribution function is given by a Gaussian PDF,

$$p(y|\beta, D^{(0)}) = \frac{1}{Z} \exp \left( -\frac{1}{2} \left( \mathcal{G}(\beta, D^{(0)}) - y \right)^\top \Sigma^{-1} \left( \mathcal{G}(\beta, D^{(0)}) - y \right) \right). \quad (11)$$

The coefficient  $Z$  is a constant of proportionality and need not be computed. The matrix  $\Sigma$  is called the error covariance matrix and is equal to  $\text{diag}(\sigma_{11}, \dots, \sigma_{JK})$ , where  $\sigma_{jk}$  is as in Eq. 8. In general the expression of the likelihood function is determined by the knowledge available about the probability distribution of the experimental errors. Since most experimental results consist of an average of many measurements, and averages typically follow a Gaussian distribution, it is commonplace to express the likelihood as a Gaussian [16].

### 3.3 Markov chain Monte Carlo

The Bayesian framework described above is nothing more than a general problem statement expressing the distribution of model parameters given noisy observations in terms of a prior and a likelihood. To solve problems posed in this framework, numerical methods are used to understand the shape of the posterior, its mean, its variance, or to compute probabilities. We choose to use a method called Markov chain Monte Carlo to understand the posterior distribution on  $\beta$  and  $D^{(0)}$ .

Given a general probability distribution function  $p(x)$ , Markov chain Monte Carlo (MCMC) methods produce a sequence of samples that, when plotted via a histogram, approximate  $p(x)$ . In the limit of infinite samples, the approximate distribution converges to  $p$ . Most MCMC methods rely on the Metropolis-Hastings algorithm [21].

In our case  $p$  will be the posterior distribution function in the Bayesian framework (Eq. 9) and the MCMC will result in  $N$  samples,  $\{\beta_n, D_n^{(0)}\}_{n=1}^N$ , from the posterior distribution. One can compute the mean and variance of these samples to obtain an estimate with associated uncertainty. There are many variants of this algorithm each with their own benefits and drawbacks [6, 28, 1, 26, 27]. Initialising the Markov chain is also a challenging task, since a bad initial choice can be detrimental to sample quality [18, 24]. Typically, the initial condition is a guess provided by the user and therefore may lie in a low probability region with respect to the posterior distribution. In what follows, we maximise the posterior distribution to find a MAP (maximum a posteriori) point before starting the sampler. This ensures the Markov chain starts in stationarity.

Samples in this work were generated by the delayed-rejection adaptive-Metropolis MCMC algorithm (DRAM) [12] implemented in the QUESO library [25]. QUESO is a C++ library for quantifying uncertainty in Bayesian calibration problems. It supports large-scale models that use many processors, and utilises multi-core architectures to provide high-quality samples from probability distributions using Markov chain Monte Carlo. QUESO is free and open source software, available at <http://libqueso.com>.

## 4 Results

### 4.1 Data available

The data available for the diffusivity of the  $(\text{CH}_4, \text{N}_2)$  couple are the measurements and fits of [32] and the fits of [20].

It is clear that the prior distribution influences the resulting posterior distribution (see Eq. 9). We chose a minimally informed prior distribution, obtained by solving the maximum entropy principle. We assert that  $D^{(0)}$  is real and non-negative and  $\beta$  is real. This provides uninformative uniform priors  $D^{(0)} \sim U[0, \infty]$  and  $\beta \sim U[-\infty, \infty]$ .

$T$ (K)	$D_{\text{CH}_4, \text{N}_2}$ ( $\text{cm}^2 \text{s}^{-1}$ )	Relative error (%)
313.7	0.242	4
314.9	0.250	4
375.2	0.353	4
474.7	0.542	4
481.0	0.539	4
573.5	0.720	4
671.1	0.919	4

Table 1: Data from [32] used for the calibration.

[32] performed their measurements with pressure  $P_{(\text{Wa})} = 1 \text{ atm}$ . Using the ideal gas assumption, we have the following conversion equations,

$$\begin{aligned} \beta_{(\text{Ma})} &= \beta_{(\text{Wa})}, \\ D^{(0)} &= A_{(\text{Wa})} \frac{P_{(\text{Wa})}}{P^0} \left( \frac{T^0}{T_{\text{ref}}} \right)^{\beta_{(\text{Wa})}}. \end{aligned} \quad (12)$$

where  $A_{(\text{Wa})}$  and  $\beta_{(\text{Wa})}$  are the [32] model parameters,  $\beta_{(\text{Ma})}$  and  $D^{(0)}$  the [20] model parameters.

## 4.2 Modeling conditions

The temperature profile and medium species density conditions under which our calibration and uncertainty propagation have been performed were chosen to match those derived from data obtained during Cassini's 40th flyby of Titan, commonly referred to as T40. The temperature profile, shown in Fig. 2, is based on fits of diffusion models to density measurements made by instruments on the Cassini orbiter and the Huygens lander as described in

## 4.3 Calibration and convergence of calibrated quantities

Fig. 3 shows  $10^4$  samples from the posterior distribution of  $D^{(0)}$ . To assess how adequately a finite number of samples represent the posterior distribution, one typically investigates the convergence of the moments of the sampled variables with respect to sampler iteration. The moments we are interested in are the mean and variance of  $D^{(0)}|_y$  and  $\beta|_y$  (Fig. 4).

Fig. 4 shows the mean ( $\mathbb{E}$ ) and variance ( $\text{Var}$ ) have converged. As expected, the mean converges faster than the variance.

The marginal distributions are approximated by histograms and Gaussian fits in Fig. 5. The distributions contain the values of the previous fits of [32] and [20] (Tab. 2).

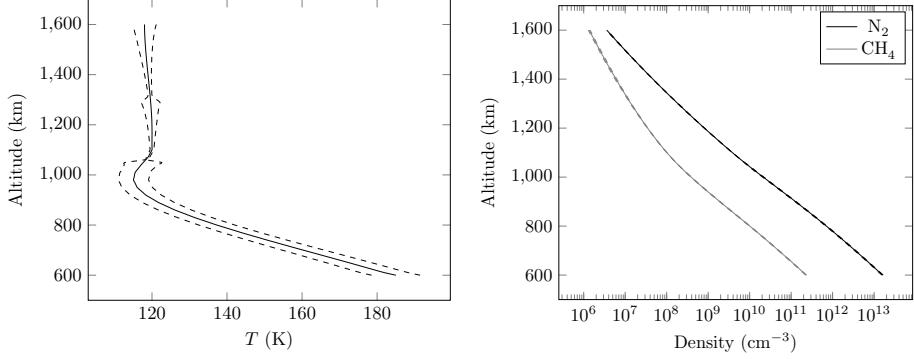


Figure 2: Estimated temperature profile of the T40 flyby (left panel) and densities of  $\text{N}_2$  and  $\text{CH}_4$  for this flyby (right panel). The uncertainties are given at a  $1 - \sigma$  level.

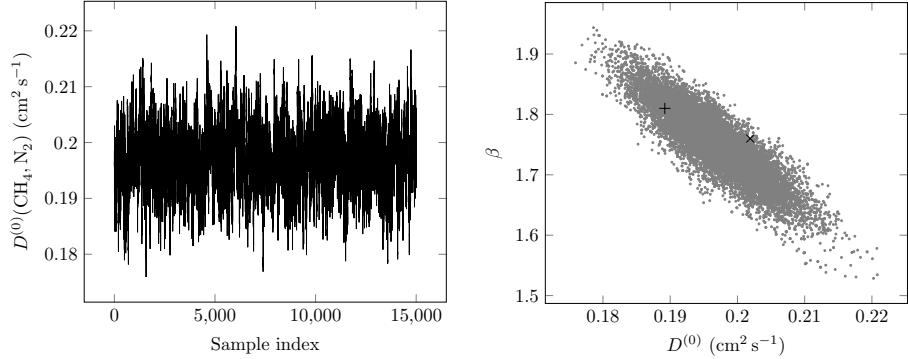


Figure 3: Left: first  $10^4$  output sample draws for  $D^{(0)}$  parameters; right:  $(\text{CH}_4, \text{N}_2)$  bimolecular diffusivity calibration. The starting point is determined by a maximization of the likelihood function. The rank correlation is  $-0.85$ . Black cross: [32] fitted values, black plus: [20] fitted values.

Model	$D^{(0)}$ ( $\text{cm}^2 \text{s}^{-1}$ )	$\beta$
Wakeham fit	0.2019	1.76
Massman fit	0.1892	1.81
This work	$0.197 \pm 0.006$	$1.75 \pm 0.05$

Table 2: Comparison of the parameters' value. The corresponding Massman parameters are given for the Wakeham value using the conversion equation 12

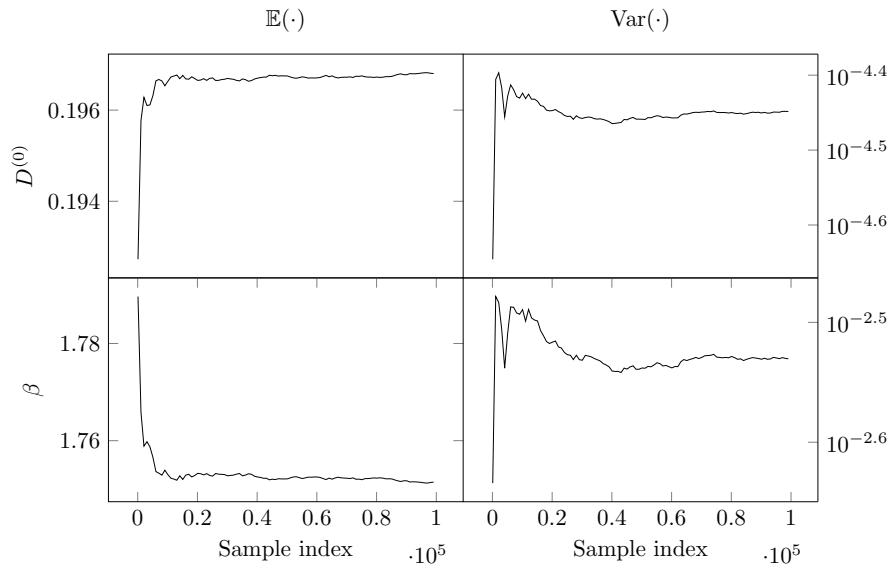


Figure 4: Means and variances on the generated samples.

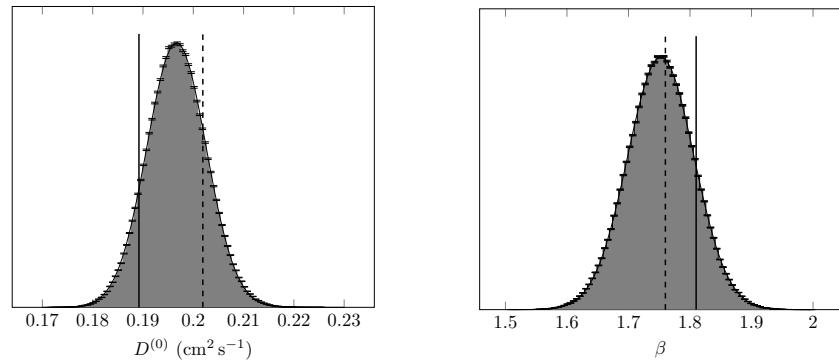


Figure 5: Posterior histograms. Vertical dashed line: [32] fitted values, vertical solid line: [20] fitted values.

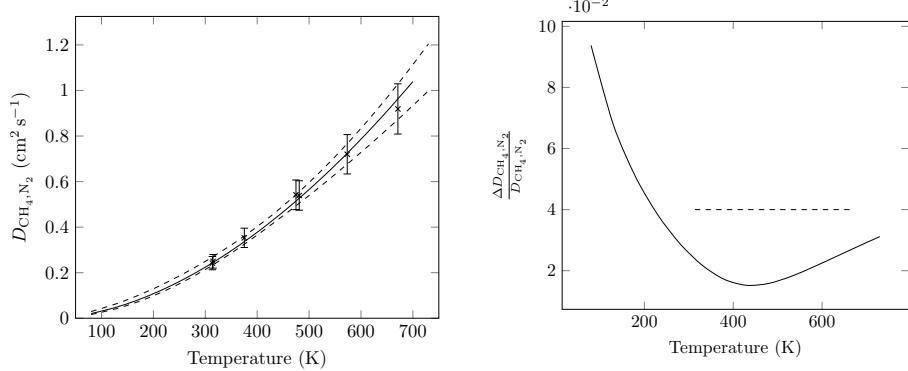


Figure 6: Estimated bimolecular diffusivity of  $(\text{CH}_4, \text{N}_2)$  at 1 atm. The dashed lines show the  $3 - \sigma$  interval.

Figure 7: Relative uncertainty of the bimolecular diffusivity of the couple  $(\text{CH}_4, \text{N}_2)$  at 1 atm. Solid: estimated, dashed: measurements.

#### 4.4 Propagation and convergence of propagated quantities

The propagated values of  $D_{\text{CH}_4, \text{N}_2}$  are given in Fig. 6, and its relative uncertainty is shown in Fig. 7. The mixture diffusion coefficient of  $\text{N}_2$  and its uncertainty are given in Fig. 8 and Fig. 9. Fig. 10 shows the convergence of the relative uncertainty of the molecular diffusion coefficient of  $\text{N}_2$ , which was found to be the most difficult to converge, requiring as many as  $2 \times 10^6$  draws.

### 5 Discussion

We have calibrated the Massman parameters for the bimolecular diffusion coefficient for the  $(\text{CH}_4, \text{N}_2)$  pair using an uninformed prior. The posterior distribution is given by the product of the likelihood and the prior (Eq. 9). Therefore the posterior is informed by data through the likelihood and previously known knowledge through the prior distribution. The choice of the prior ensures a maximal effect of the information of the data in the Shannon entropy sense.

As figures 3 and 5 show, the previous fits from [32] and [20] are but one point in the posterior space. The cross/dashed lines are [32] values, the plus/solid lines are [20] values. Although the fits accurately reproduce the diffusivity at Titan's temperatures, all information about the uncertainty is lost.

We observe a negative correlation between the relative uncertainty of the molecular diffusion of  $\text{N}_2$  and the temperature (Fig. 2 and 9). A sensitivity analysis of the Massman model (Eq. 13) yields a  $\ln(T/T^0)$  term with respect to the  $\beta$  parameter that can explain this behavior. Titan's temperatures being lower than the reference temperature of the Massman model ( $T^0 = 273.15$  K) means this log-term will diverge to  $-\infty$ , therefore increasing the uncertainty

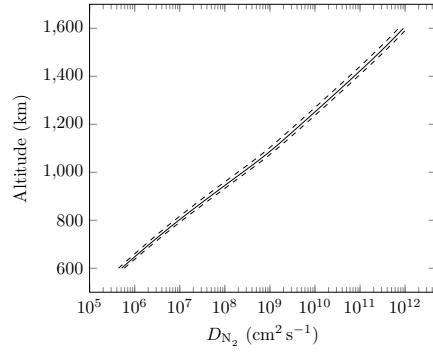


Figure 8: Wilke diffusivity of  $\text{N}_2$ , the  $3 - \sigma$  interval is shown with the dash lines.

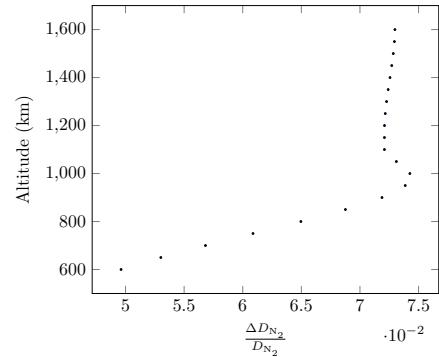


Figure 9: Wilke diffusivity relative uncertainty of  $\text{N}_2$ .

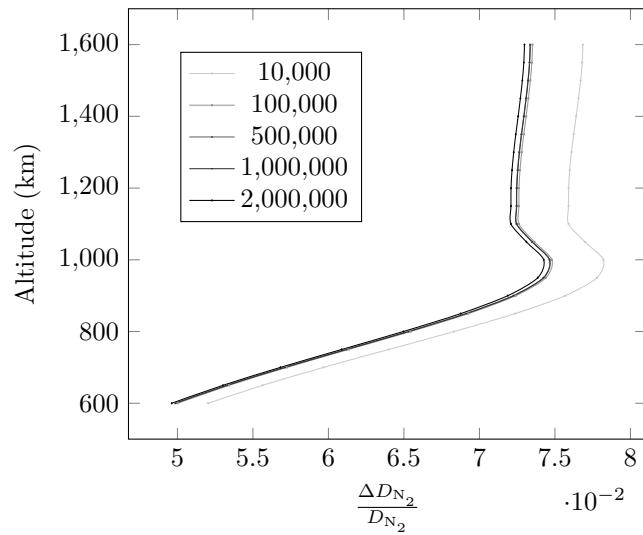


Figure 10: Convergence of the relative uncertainty of the molecular diffusion coefficient of  $\text{N}_2$ .

of the bimolecular diffusion coefficient (Fig. 7) with respect to  $\beta$ . A sensitivity of the molecular diffusion coefficient with respect to the  $\beta$  parameter could thus explain the negative correlation between the molecular diffusion relative uncertainty and the temperature.

$$\begin{aligned}\frac{\partial D_{s,m}}{\partial D^{(0)}} &= \frac{P^0}{P} \left( \frac{T}{T^0} \right)^\beta, \\ \frac{\partial D_{s,m}}{\partial \beta} &= D^{(0)} \frac{P^0}{P} \left( \frac{T}{T^0} \right)^\beta \ln \left( \frac{T}{T^0} \right).\end{aligned}\quad (13)$$

One important consequence is that at lower temperatures, investigations to reduce uncertainties on molecular diffusion should focus on the  $\beta$  parameter, typically by characterizing its value at these temperatures.

In general, the uncertainties in molecular diffusion are small compared to other sources of uncertainty, which can be orders of magnitude larger [13, 4]. However, a propagation of the uncertainties in binary molecular diffusion to the calculation of the methane mixing ratio as a function of altitude, illustrated in Fig. 11, demonstrates that the uncertainty is nontrivial above 1200 km. This methane mixing ratio profile is determined using an eddy coefficient that was optimized for the methane profile and temperature and density profiles based on the T40 flyby as was done in [19]. Also shown are the methane mixing ratios measured by the Cassini Ion Neutral Mass Spectrometer [31] during the T40 flyby. Uncertainties in molecular diffusion coefficients have the greatest impact at the highest altitudes, because it is at these altitudes where molecular diffusion has a dominant effect compared to eddy diffusion. This has important implications for studies of Titan's atmosphere that use models to evaluate escape rates based on mixing ratios modeled above 1200 km [7], suggesting that uncertainties in molecular diffusion need to be taken into account when making conclusions about measured altitude profiles.

For the common situation where no measurement data is available one has to resort to other theoretical or model calculations. In the case of missing laboratory gas mixture data, diffusion coefficients may be estimated by a transport theory calculation. The Bayesian framework can also handle this situation; the likelihood distribution will, instead of acting on lab measurement data, act on transport theory calculation output. It should thus capture the uncertainties and inadequacies associated to the theoretical or model calculations. This requires a careful analysis of the corresponding model that is beyond the scope of this work.

In the situations where transport theory calculations yield poor diffusion coefficient estimates, it is likely that the uncertainty in the diffusion coefficients would be inflated. Access to lab measurements offers the possibility of providing favourable uncertainty calculations but investigating the effect of theoretical calculations on diffusion coefficient uncertainty is a topic of future work.

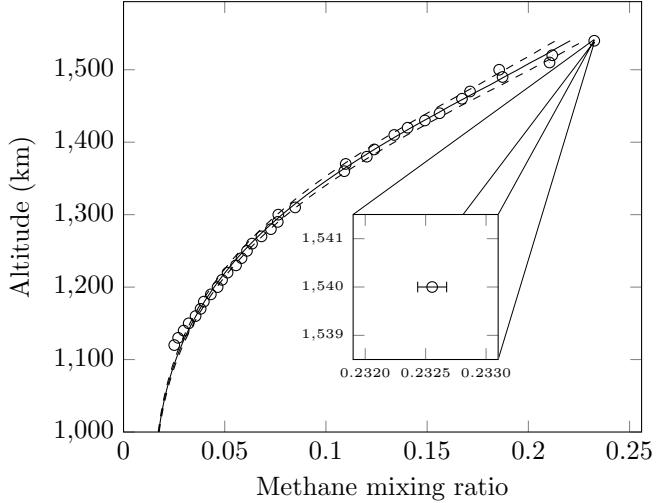


Figure 11: Propagation of the uncertainties in molecular diffusion, roughly  $\pm 7.5\%$ , to uncertainties in modeled methane mixing ratio. The data points are the measurements derived from CASSINI; note that the statistical uncertainties from the INMS are orders of magnitude smaller.

## 6 Acknowledgements

We acknowledge the financial help of NASA OPR program with the grant number NNH12ZDA001N.

## References

- [1] A Beskos, G O Roberts, and A M Stuart. Optimal scalings for local Metropolis-Hastings chains on nonproduct targets in high dimensions. *The Annals of Applied Probability*, 19(3):863–898, June 2009.
- [2] N. Carrasco, C. Alcaraz, O. Dutuit, S. Plessis, R. Thissen, V. Vuitton, R. Yelle, and P. Pernot. Sensitivity of a Titan ionospheric model to the ion-molecule reaction parameters. *Planetary and Space Science*, 56(12):1644–1657, 2008. Surfaces and Atmospheres of the Outer Planets, their Satellites and Ring Systems, Part IV, Meetings held in 2007: EGU: PS3.0 & PS3.1; IUGG/IAMAS:JMS12 & JMS13; AOGS: PS09 & PS11; EPSC2: AO4 or PM1.
- [3] N. Carrasco, O. Dutuit, R. Thissen, M. Banaszkiewicz, and P. Pernot. Uncertainty analysis of bimolecular reactions in Titan ionosphere chemistry model. *Planetary and Space Science*, 55:141–157, 2007.

- [4] N. Carrasco, E. Hébrard, M. Banaszkiewicz, M. Dobrijevic, and P. Pernot. Influence of neutral transport on ion chemistry uncertainties in Titan ionosphere. *Icarus*, 192(2):519–526, 2007.
- [5] N. Carrasco and P. Pernot. Modeling of branching ratio uncertainty in chemical networks by Dirichlet distributions. *Journal of Physical Chemistry A*, 111(18):3507–3512, 2007.
- [6] S L Cotter, G O Roberts, A M Stuart, and D White. MCMC Methods for functions: Modifying old algorithms to make them faster. 2012.
- [7] J. Cui, R. V. Yelle, D. F. Strobel, I. C. F. Müller-Wodarg, D. S. Snowden, T. T. Koskinen, and M. Galand. The CH<sub>4</sub> structure in Titan’s upper atmosphere revisited. *Journal of Geophysical Research: Planets*, 117(E11):n/a–n/a, 2012.
- [8] Virginie De La Haye. *Corona Formation and Heating Efficiencies in Titan’s Upper Atmosphere: Construction of a coupled Ion, Neutral and Thermal Structure Model To Interpret the First INMS Cassini Data*. PhD thesis, The University of Michigan, 2005.
- [9] M. Dobrijevic, N. Carrasco, E. Hébrard, and P. Pernot. Epistemic bimodality and kinetic hypersensitivity in photochemical models of Titan’s atmosphere. *Planetary and Space Science*, 56(12):1630–1643, 2008. Surfaces and Atmospheres of the Outer Planets, their Satellites and Ring Systems, Part IV, Meetings held in 2007: EGU: PS3.0 & PS3.1; IUGG/IAMAS:JMS12 & JMS13; AOGS: PS09 & PS11; EPSC2: AO4 or PM1.
- [10] G.J. Erickson, J.T. Rychert, and C.R. Smith, editors. *Maximum Entropy and Bayesian Methods*. Kluwer Academic, Dordrecht, 1998.
- [11] B. Gans, Z. Peng, N. Carrasco, D. Gauyacq, S. Lebonnois, and P. Pernot. Impact of a new wavelength-dependent representation of methane photolysis branching ratios on the modeling of Titans atmospheric photochemistry. *Icarus*, 223(1):330–343, 2013.
- [12] Heikki Haario, Marko Laine, Antonietta Mira, and Eero Saksman. DRAM: Efficient adaptive MCMC. *Statistics and Computing*, 16(4):339–354, December 2006.
- [13] E. Hébrard, M. Dobrijevic, Y. Bénilan, and F. Raulin. Photochemical kinetics uncertainties in modeling Titan’s atmosphere: a review. *Journal of Photochemistry and Photobiology A: Chemistry*, 7:211–230, 2006.
- [14] E. Hébrard, M. Dobrijevic, Y. Bénilan, and F. Raulin. Photochemical kinetics uncertainties in modeling Titan’s atmosphere: First consequences. *Planetary and Space Science*, 55:1470–1489, 2007.

- [15] E. Hébrard, M. Dobrijevic, P. Pernot, N. Carrasco, A. Bergeat, K. M. Hickson, A. Canosa, S. D. Le Picard, and I. R. Sims. How measurements of rate coefficients at low temperature increase the predictivity of photochemical models of Titan’s atmosphere. *The Journal of Physical Chemistry A*, 113(42):11227–11237, October 2009.
- [16] JCGM. Evaluation of measurement data - Guide to the expression of uncertainty in measurement (GUM). Technical Report 100, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, 2008. GUM 1995 with minor corrections.
- [17] D. Jenkinson. The elicitation of probabilities - A review of the statistical literature. Beep working paper, BEEP working paper, University of Sheffield, 2005.
- [18] Mary Kathryn, Bradley P Carlin, Mary Kathryn Cowles, and Bradley P Carlin. Markov Chain Monte Carlo Convergence Diagnostics : A Comparative Review. 91(434):883–904, 2014.
- [19] Kathleen E. Mandt, David A. Gell, Mark Perry, J. Hunter Waite, Frank A. Crary, David Young, Brian A. Magee, Joseph H. Westlake, Thomas Cravens, Wayne Kasprzak, Greg Miller, Jan-Erik Wahlund, Karin Ågren, Niklas J. T. Edberg, Alan N. Heays, Brenton R. Lewis, Stephen T. Gibson, V. de la Haye, and Mao-Chang Liang. Ion densities and composition of Titan’s upper atmosphere derived from the Cassini Ion Neutral Mass Spectrometer: Analysis methods and comparison of measured ion densities to photochemical model simulations. *Journal of Geophysical Research: Planets*, 117(E10), 2012.
- [20] W.J. Massman. A review of the molecular diffusivities of  $\text{H}_2\text{O}$ ,  $\text{CO}_2$ ,  $\text{CH}_4$ ,  $\text{CO}$ ,  $\text{O}_3$ ,  $\text{SO}_2$ ,  $\text{NH}_3$ ,  $\text{N}_2\text{O}$ ,  $\text{NO}$ , and  $\text{NO}_2$  in air,  $\text{O}_2$  and  $\text{N}_2$  near STP. *Atmospheric Environment*, 32(6):1111–1127, 1998.
- [21] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [22] Zhe Peng, Michel Dobrijevic, Eric Hebrard, Nathalie Carrasco, and Pascal Pernot. Photochemical modeling of Titan atmosphere at the “10 percent uncertainty horizon”. *Faraday Discuss.*, 147:137–143, 2010.
- [23] Sylvain Plessis, Nathalie Carrasco, and Pascal Pernot. Knowledge-based probabilistic representations of branching ratios in chemical networks: the case of dissociative recombination. *Journal of Chem. Phys.*, 133(13):134110–134131, 2010.
- [24] Martyn Plummer, Nicky Best, Kate Cowles, and Karen Vines. CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6(1):7–11, 2006.

[25] Ernesto E. Prudencio and Karl W. Schulz. The Parallel C++ Statistical Library ‘QUESO’: Quantification of Uncertainty for Estimation, Simulation and Optimization. In Michael Alexander, Pasqua D’Ambra, Adam Belloum, George Bosilca, Mario Cannataro, Marco Danelutto, Beniamino Di Martino, Michael Gerndt, Emmanuel Jeannot, Raymond Namyst, Jean Roman, Stephen L. Scott, Jesper Larsson Traff, Geoffroy Valle, and Josef Weidendorfer, editors, *Euro-Par 2011: Parallel Processing Workshops*, volume 7155 of *Lecture Notes in Computer Science*, pages 398–407. Springer Berlin Heidelberg, 2012.

[26] G O Roberts. Weak convergence and optimal scaling of random walk Metropolis Algorithms. *Annals of Applied Probability*, 7(1):110–120, 1997.

[27] G O Roberts and J S Rosenthal. Optimal scaling of discrete approximations to Langevin diffusions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(1):255–268, February 1998.

[28] G O Roberts and J S Rosenthal. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical Science*, 16(4):351–367, November 2001.

[29] Hang K. Ryu. Maximum entropy estimation of density and regression functions. *Journal of Econometrics*, 56(3):397–440, 1993.

[30] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423,623–656, 1948.

[31] J.H. Waite, W.S. Lewis, W.T. Kasprzak, V.G. Anicich, B.P. Block, T.E. Cravens, G.G. Fletcher, W.-H. Ip, J.G. Luhmann, R.L. Mcnutt, H.B. Niemann, J.K. Parejko, J.E. Richards, R.L. Thorpe, E.M. Walter, and R.V. Yelle. The Cassini Ion and Neutral Mass Spectrometer (INMS) Investigation. *Space Science Reviews*, 114(1):113–231, 2004.

[32] W. A. Wakeham and D. H. Slater. Diffusion coefficients for n-alkanes in binary gaseous mixtures with nitrogen. *Journal of Physics B: Atomic and Molecular Physics*, 6(5):886–896, 1973.

[33] C. R. Wilke. A Viscosity Equation for Gas Mixtures. *The Journal of Chemical Physics*, 18(4):517, 1950.

[34] E. H. Wilson. *Investigations into the photochemistry of the current and primordial atmosphere of Titan*. PhD thesis, University of Michigan, University of Michigan, 2002.

[35] E.H. Wilson and S.K. Atreya. Current state of modeling the photochemistry of Titan’s mutually dependent atmosphere and ionosphere. *Journal of Geophysical Research*, 109:E06002, 2004.